

## 자기구성지도 기반 방법을 이용한 이상 탐지 Novelty Detection using SOM-based Methods

이형주, 조성준

E-mail: {impatton, zoon}@snu.ac.kr

서울대학교 공과대학 산업공학과

서울시 관악구 신림동 산 56-1, 151-744

**Abstract** Novelty detection involves identifying novel patterns. They are not usually available during training. Even if they are, the data quantity imbalance leads to a low classification accuracy when a supervised learning scheme is employed. Thus, an unsupervised learning scheme is often employed ignoring those few novel patterns. In this paper, we propose two ways to make use of the few available novel patterns. First, a scheme to determine local thresholds for the Self Organizing Map boundary is proposed. Second, a modification of the Learning Vector Quantization learning rule is proposed so that allows one to keep codebook vectors as far from novel patterns as possible. Experimental results are quite promising.

**Keywords** Novelty detection, Novel data, Closed boundary, Codebook methods, Self-organizing map, Learning vector quantization

### 1. 서론

일반적인 이진 분류 문제에서는 두 가지 클래스를 모두 학습한 후, 새로운 입력 패턴이 주어지면 두 가지 클래스 중 하나로 구분한다. 그러나 실제 문제에서는 한 쪽 클래스에 속하는 패턴의 수가 극히 적거나 수집이 어려운 경우가 있다. 예를 들어, 지폐 인증 문제에서 ‘위폐’라는 패턴을 사전에 미리 구하여 이를 이용한 분류기를 구축한다는 것은 현실

적으로 불가능하다. 따라서, 이와 같은 경우에 이상 탐지 방법론을 적용하여 분류 문제를 해결할 수 있다[1]. 즉, 다수 클래스를 “정상” 클래스로 두고 정상 데이터의 특성을 학습한 후, 정상 데이터와 매우 다른 특성을 보이는 “이상” 패턴을 감지하는 작업을 하는 것이다. 이를 기하학적으로 해석하면, 정상 패턴 주위로 닫힌 경계(closed boundary)를 생성하는 것으로 볼 수 있다[11].

이상 탐지 작업을 위하여 다양한 기법들이 적용되어 왔다[7-9]. 대부분의 기법들은 정상 데이터를 사용하여 학습하고, 학습이 끝난 후 threshold를 정하거나 모델의 모수를 결정하는 과정에서 이상 데이터를 사용하게 된다. 몇몇의 예[2, 3, 6, 12]를 제외하면, 학습 과정에서 이상 데이터가 존재한다고 가정하지 않기 때문이다. 그러나, 많은 경우에, 비록 소수이지만 이상 패턴들이 학습 데이터셋에 존재한다. 이상 패턴의 수가 이진 분류기를 학습하기에는 부족한 경우라도, 정상 클래스의 경계를 결정하는데 도움을 줄 수 있다. 예를 들어, 기계의 고장 탐지를 생각해 보자. 처음 얼마 동안은 기계가 고장없이 작동할 것이다. 즉, 정상 클래스의 데이터만 존재한다. 그러나, 시간이 지나면 극소수의 이상 데이터, 즉 “고장 패턴”도 수집될 것이다. 이 때, 정상 데이터만으로 시스템을 구축하는 것은 이상 데이터에 포함되어 있는 유용한 정보를 버리는 결과를 초래할 수 있다. 이상 탐지 기법의 학습 과정에서 이상 데이터를 활용하면, 더 좋은 분류 성능을 얻을 수 있다는 사실이 실험을 통하여 입증된 바 있다[2, 3, 12].

정상 클래스를 정의하는 모델이 해당 클래스의

데이터를 되도록 많이 포함하려고 하는 것을 generalization, 다른 클래스의 데이터를 가능한 한 적게 포함하려고 하는 것을 specialization이라고 한다[4]. Generalization과 specialization의 균형이 맞았을 때 좋은 분류 성능을 얻을 수 있다. 많은 이상 탐지 기법들은 하나의 클래스, 즉 정상 클래스만을 고려하기 때문에, generalization은 자연스럽게 가능하지만 specialization은 사용자가 제공한 특정한 bias에 따라서만 가능하다. 따라서, 이상 탐지 문제의 학습과정에서 이상 데이터를 사용하는 것은, generalization 뿐만 아니라 specialization도 데이터를 통해 달성하도록 돕는 과정이라고 생각할 수 있다.

본 논문에서는 자기구성지도 (self-organizing map; SOM)를 기반으로 하여, 학습 데이터에 이상 패턴이 극히 소수 존재하는 경우에 적용될 수 있는 두 가지 방법을 제안한다. 첫 번째 방법은 기존의 SOM과 동일하게 코드북 조정이 이루어지지만, threshold 결정 과정에서 두 클래스를 모두 고려할 수 있도록 개선한 SOM with local thresholds(SOM-L)이다. 두 번째 방법은, 이상 데이터를 코드북 조정 과정에 반영할 수 있는 새로운 이상 탐지 기법인 one-class LVQ(OneLVQ)이다. 이 방법에서는, 정상 패턴을 둘러싸는 경계를 정하는 과정에서 그 영역이 이상 패턴을 포함하지 않도록 한다. 학습이 끝나면, SOM-L과 마찬가지로 두 클래스의 데이터를 바탕으로 threshold를 결정하게 된다. 궁극적으로는 정상 데이터만을 사용했을 때보다 더 정확하고 정밀한 경계를 생성할 수 있을 것으로 기대된다.

다음 절에서는 제안된 두 가지 방법, SOM-L과 OneLVQ에 대하여 설명한다. 3절에서는 위의 방법들을 인공 데이터와 실제 데이터에 적용하여 다른 이상 탐지 기법과 비교한다. 4절에서 결론을 제시하고 추후 과제에 대하여 토의한다.

## 2. 제안 방법

### 2.1 이상 탐지를 위한 자기구성지도

학습 데이터  $\mathbf{X} = \{\mathbf{x}_i\}_{i=1}^N$ 가 주어져 있다고 가정해 보자. 여기에서  $\mathbf{x}_i \in \mathbf{R}^d$ 는 입력 패턴이고  $y_i \in \{+1, -1\}$ 는 클래스 레이블이다. 정상 클래스와 이상 클래스는 각각  $\mathbf{T} = \{\mathbf{x}_i | y_i = +1\}$ 와  $\mathbf{O} = \{\mathbf{x}_i | y_i = -1\}$ 로 표기한다. 일반적으로, 정상 패턴의 수가 이상 패턴의 수보다 훨씬 많다. 즉,  $(|\mathbf{T}| \gg |\mathbf{O}|)$ . 기존의 SOM[5]에서는 정상 패턴만 학습에 사용된다. SOM은  $K (\ll N)$ 개의 코드북을 생성하여 정상 데이터를 설명하게 된다.

코드북 조정이 끝나면, 입력 패턴  $\mathbf{x}$ 의 코드북 벡터  $\mathbf{m}(\mathbf{x})$ 와 각 코드북  $\mathbf{w}_k$ 의 Voronoi region  $\mathbf{S}_k$ 가 다음과 같이 정의된다.

$$\mathbf{m}(\mathbf{x}) = \mathbf{w}_k \Leftrightarrow \mathbf{x} \in \mathbf{S}_k, \quad (1)$$

$$\text{if } \|\mathbf{w}_k - \mathbf{x}\| < \|\mathbf{w}_l - \mathbf{x}\|, \forall l \neq k.$$

테스트 패턴  $\mathbf{z}$ 가 주어지면,  $\mathbf{z}$ 와  $\mathbf{m}(\mathbf{z})$  사이의 유클리드 거리  $e(\mathbf{z})$ 가 어떤 threshold보다 크면 이상으로 분류하고, 그렇지 않으면 정상으로 분류하게 된다. 기존의 SOM은 하나의 전역적(global) threshold를 사용하기 때문에, 이것을 SOM-G라고 부르기로 한다.  $\mathbf{z}$ 로부터 어느 코드북이 가장 가까운지에 관계없이 단 하나의 threshold를 기준으로 정상과 이상을 판정한다. 이상으로 판정될 정상 패턴의 비율,  $\theta$ 를 미리 정하고, 이에 따라 threshold를 결정한다. 즉,

$$e(\mathbf{z}) = \|\mathbf{z} - \mathbf{m}(\mathbf{z})\|^2 \leq r^2, \quad (2)$$

여기에서,  $\frac{1}{N} \sum_{\mathbf{x}_i \in \mathbf{T}} I\{e(\mathbf{x}_i) > r^2\} = \theta$  이고,  $I\{\}$ 는 indicator 함수이다.

### 2.2 지역적(local) thresholds의 결정

입력 패턴들이 밀집된 영역에 위치한 코드북이

있는 반면, 드물게(sparsely) 분포된 입력 패턴들의 영역에 할당되는 코드북도 있다. (2)와 같이 모든 코드북들에 대하여 동일한 threshold를 취하게 되면, 입력 패턴의 분포에 비해 너무 넓은 영역을 책임지게 되거나 너무 많은 정상 패턴을 경계 바깥에 두는 일이 생길 수 있다. 따라서, 코드북마다 각기 다른 threshold를 설정하는 것이 바람직하다.

여기에서는 support vector data description(SVDD) [12]와 유사한 threshold 결정 방법을 제안한다. 코드북 조정이 끝나면, 각 패턴은 해당 Voronoi 영역  $S_k$ 에 포함되고, 각 Voronoi 영역은 코드북 벡터  $w_k$ 를 가진다. 여기에서는 최소한의 반지름을 가지면서, 되도록이면 많은 정상 패턴을 둘러싸고 가능한 한 이상 패턴은 포함하지 않는,  $w_k$ 를 중심으로 하는 hypersphere를 Voronoi 영역마다 구하게 된다. 그러면 모든 Voronoi 영역  $S_k$ ,  $\forall k$ 에 대하여, 다음의 “최적화” 문제를 생각할 수 있다.

$$\min \bar{E}(r_k) = r_k^2 + C_1 \sum_{y_i=1} \varepsilon_i + C_2 \sum_{y_i=-1} \xi_i, \quad (3)$$

$$\text{s.t. } \begin{cases} \|x_i - w_k\|^2 \leq r_k^2 + \varepsilon_i, & \text{if } x_i \in T_k \\ \|x_i - w_k\|^2 \geq r_k^2 - \xi_i, & \text{if } x_i \in O_k \end{cases} \quad (4)$$

$$\varepsilon_i, \xi_i \geq 0, \forall i.$$

여기에서  $T_k = T \cap S_k$  이고  $O_k = O \cap S_k$ 이다. 반지름이 큰 hypersphere를 그리게 되면, 많은 정상 패턴을 포함할 수 있지만, 영역이 넓어진만큼 false acceptance의 가능성이 높아진다. 반대로, 반지름이 너무 작으면, 이상 패턴을 받아들이지 않을 수 있지만 정상 패턴도 너무 많이 기각할 수 있다. 그래서, 양의 상수  $C_1$ 과  $C_2$ 를 도입하여 반지름과 두 오류 사이의 trade-off를 고려하게 된다. (3)의 해  $r_k^*$ 는 전수 조사(exhaustive search)를 통해서,  $|T_k|$  회의 계산만으로 구할 수 있다. 결국, 밀집된 영역에는 작은 threshold를, 넓게 퍼져 있는 영역에는 큰 threshold를 찾게 된다.

$$(r_k^*)^2 = e(x_u) \quad , \quad \text{where } u = \arg \min_{x_i \in T_k} \bar{E}(e(x_i)) \quad (5)$$

각 코드북의 최적 threshold는 독립적으로 구할 수 있다. 테스트 패턴  $z$ 는, (6)을 만족하면 정상으로 ( $z \in T$ ) 판정되고, 그렇지 않으면 이상으로( $z \in O$ ) 판정된다.

$$\|z - w_q\|^2 \leq (r_q^*)^2 \quad (6)$$

SOM-L에서는 기존의 SOM을 학습하지만, 새로운 테스트 패턴을 분류할 때에는 (6)과 같은 지역적 thresholds가 사용된다.

### 2.3 수정된 학습 규칙

여기에서 제안하는 OneLVQ는 기존 LVQ의 학습 규칙을 수정한 것이다. LVQ는 모든 클래스에 코드북을 할당하고, 패턴이 주어졌을 때에는 가장 가까운 코드북이 속한 클래스로 분류한다. 반면, OneLVQ는 단 하나의 클래스, 즉 정상 클래스에만 코드북을 할당한다. 그런 이유로 one-class LVQ라고 부르게 되었다.

학습 데이터셋에 이상 패턴이 포함되어 있을 때, 수정된 오류 함수는 다음과 같이 쓸 수 있다.

$$\begin{aligned} E &= \int y \|x - m(x)\|^2 p(x)p(y) dx dy \\ &\cong \frac{1}{N} \sum_i y_i \|x_i - m(x_i)\|^2 \\ &= \frac{1}{N} \sum_k \left[ \sum_{x_i \in T_k} \|x_i - w_k\|^2 - \sum_{x_i \in O_k} \|x_i - w_k\|^2 \right]. \end{aligned} \quad (7)$$

이 오류 함수를 최소화하면 정상 패턴에는 최대한 가깝게, 이상 패턴에는 최대한 멀게 코드북이 위치할 것이다. 따라서, 기존 LVQ의 학습 규칙과는 다른 학습 규칙이 나오게 된다.

새로운 입력 패턴  $x$ 가 주어지면,

$$w_k \leftarrow \begin{cases} w_k & x \notin S_k \\ w_k + \eta(x - w_k) & x \in T_k \\ w_k - \eta(x - w_k) & x \in O_k \end{cases} \quad (8)$$

이 규칙에 따르면,  $x$ 가 Voronoi 영역  $S_k$ 에 속하지 않으면,  $w_k$ 의 위치는 바뀌지 않는다. 반대로, 주어진 입력 패턴이  $S_k$ 에 속할 때에는,  $x$ 가 정상 클래스에 속하면 코드북  $w_k$ 가  $x$  쪽으로 이동하고,  $x$

가 이상 클래스에 속하면  $w_k$ 가  $x$ 의 반대쪽으로 이동한다. 다시 말해서, 정상 패턴들은 코드북 벡터를 “당기고”, 이상 패턴들은 코드북 벡터를 “밀어내게” 된다.

OneLVQ는 기존의 SOM을 바탕으로 초기화되고, (8)의 학습 규칙에 따라 학습된다. 분류 과정에서는 SOM-L과 마찬가지로 (6)의 지역적 threshold를 사용하게 된다.

### 3. 실험 결과

#### 3.1 인공 데이터셋

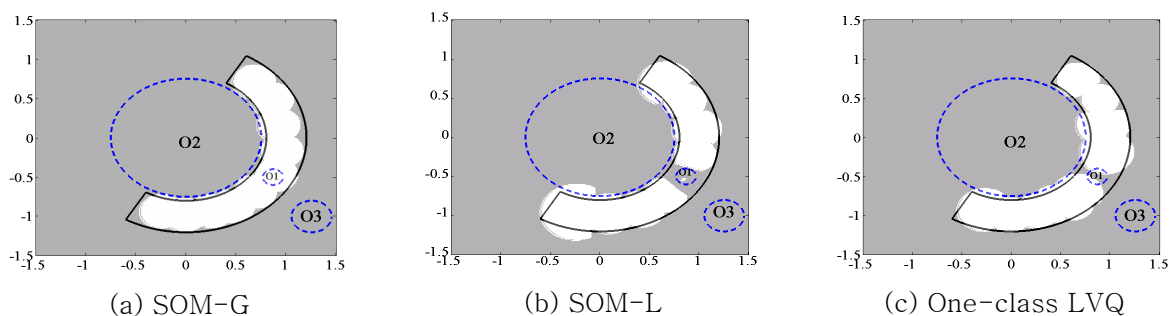
먼저, SOM-L과 OneLVQ의 이상 탐지 성능을 직관적으로 보여주기 위하여 간단한 2차원 인공 데이터셋에 대하여 실험을 수행하였다. 말굽 형태의 데이터를 3개의 이상 탐지 기법, SOM-G, SOM-L, OneLVQ로 학습하였다. 그 분류 경계들이 [그림 1]에 표시되어 있다. 정상과 이상 영역의 실제 경계는 각각 실선과 점선으로 표시되어 있다. 이상 탐지 문제에서의 중요한 가정은 학습 데이터에 주어진 이상 패턴의 영역 이외에서도 이상 패턴들이 나타날 수 있다는 것이다. 따라서, O1 영역에서의 이상 패턴들은 학습 데이터에 주어져 있지만, O2와 O3 영역에서는 학습 데이터에는 이상 패턴이 주어져 있지 않고, 테스트 단계에서만 이상 패턴이 발생한다고 가정했다.

각 모델은 밝게 표시된 부분을 정상으로, 어둡게 표시된 부분을 이상으로 분류하였다. 모델들의

모수는 실험적으로 결정되었다. SOM-G는 정상 영역 바깥에 있는 O2, O3 영역은 이상으로 구분해 내었지만, 정상 영역 내부에 있는 O1은 구분해내지 못했다. 그 이유는 이상 영역인 O1을 학습에서 사용할 수 없었기 때문이다. SOM-L의 경우에는 O1 영역을 정상으로 구분하였지만, 정상 영역의 많은 부분을 이상으로 오분류하는 문제점을 보였다. 한편, OneLVQ는 O1, O2, and O3를 모두 이상으로 분류하여, 코드북 조정 과정에서 이상 데이터를 활용하는 것이 그렇지 않은 것보다 더 바람직함을 보였다.

#### 3.2 Rätsch의 벤치마크 데이터셋

[10]에서 쓰인 벤치마크 데이터셋 중에서 Banana, Breast Cancer, Diabetes, German, Heart, Titanic의 6개 데이터셋에 대하여 실험하였다. <http://ida.first.fraunhofer.de/~raetsch/>에서 위의 데이터셋들을 얻을 수 있다. 각 dataset은 학습과 테스트셋 각각 100개씩의 split으로 이루어져 있고, 이 실험에서는 각 데이터마다 30개씩의 splits를 사용하였다. 위의 6개 데이터셋은 모두 이진 분류 문제를 위한 것이기 때문에, 이상 탐지 방법을 사용하기 위해 데이터를 변형시켰다. 우선, 다수 클래스를 정상 클래스로, 소수 클래스를 이상 클래스로 정의하였다. 그리고, 이상 데이터가 드문 경우를 가정하여, 이상 데이터 중에서 랜덤 샘플링하여 학습 데이터에서 이상 패턴의 수가 정상 패턴의 10%가 되도록 하였다.



[그림 1] 말굽 형태의 데이터에 대한 3개 이상 탐지 기법의 분류 경계

분류 성능 지표로서 area under receiver operating characteristic (AUROC)를 사용하였다. 이 값은 false rejection rate(FRR)을 1%에서 50%까지 변화시키면서 true rejection rate(TRR)을 그린 ROC 곡선으로부터 계산되었다. [표 1]은 30개 splits에 대한 반복 실험 후, AUROC의 평균값을 정리한 것이다. 6개의 데이터셋 중 Banana를 제외한 5개에서 OneLVQ가 가장 높은 AUROC를 보인다. 그 중에서도, Diabetes, German, and Titanic에 대하여는 유의수준 5%에서 통계적으로 유의한 차이로 다른 모델들보다 좋은 AUROC 값을 가졌다. SOM-L은 대부분의 데이터에서 OneLVQ에 이어 두 번째로 높은 성능을 보였다. OneLVQ와 SOM-L의 차이점은, OneLVQ는 코드북 조정 과정에서 이상 클래스를 반영시켰고 SOM-L은 그렇지 않았다는 점이다. 따라서, 이상 데이터를 학습에 사용한 것이 OneLVQ와 SOM-L의 성능 차이의 원인이라고 생각해 볼 수 있다. 한편, SOM-L은 SOM-G보다 대체로 높은 성능을 보이는 것으로 보아 제안된 threshold 결정 방법이 효과가 있는 것으로 해석할 수 있다. 그리고, one-class support vector machine(OneSVL)과 SVDD의 AUROC는 정규분포 추정 방법(Gauss)과 비슷한 것으로 보아, 두 support vector 기반 방법들의 분류 경계는 hypersphere와 거의 동일한 것으로 추측할 수 있다. Parzen 분포 추정 방법은 2차원인 Banana에서는 좋은 성능을 보였지만, 다른 데이터셋에서는 Gauss와

크게 다르지 않았다. 자기연상 신경망 (auto

o-associative neural network; AANN)은 대체로 낮은 AUROC를 나타내었다.

### 3.3 타자 패턴 데이터셋

실제 문제로서, 이상 탐지 기법들을 타자 패턴 데이터셋에 적용하였다. 이 문제의 목적은 컴퓨터 사용자의 암호 타자 패턴의 특성을 파악하여 잠재적인 침입자의 타자 패턴을 감지해 내는 것으로, 궁극적으로는 컴퓨터 시스템을 침입으로부터 보호하는 것이다. 21명의 사용자가 각자 자신의 암호를 타자해서 정상 클래스의 데이터를 생성하고, 15명의 “침입자”가 21명 사용자의 암호를 타자해서 가상의 침입 상황의 이상 데이터를 생성하였다. 그래서, 21명의 사용자에게 대하여 21개의 데이터셋이 생성되었다. 각 사용자의 암호에 대하여, 76~388개의 정상 학습 패턴과 75개의 정상 테스트 패턴, 그리고 75개의 이상 패턴이 수집되었다. 이 데이터셋에 대한 더 자세한 설명은 [13]에 기술되어 있다. 정상 학습 패턴 중에서 50개, 이상 패턴 중에서 5개를 랜덤 샘플링하여 학습 데이터셋을 구성하였다. 그리고, 75개의 정상 테스트 패턴과 나머지 70개의 이상 패턴으로 테스트 데이터셋이 구성되었다. 10개의 서로 다른 학습 및 테스트 데이터셋이 랜덤 샘플링을 통하여 구성되었다. 그리고, 주성분 분석을 통하여 80%의 variance를 설명하도록 변수의 수를 대략 반 정도로 감소시켰다.

[표 2]에 7개 이상 탐지 기법들의 21개 암호에 대한 false rejection rates(FRR)와 false acc

[표 1] 6개 벤치마크 데이터셋에 대한 AUROC (%)

Dataset	Dim	Gauss	Parzen	OneLVQ	SOM-L	SOM-G	AANN	OneSVM	SVDD
Banana	2	16.25	<b>42.73*</b>	41.85	40.93	41.06	39.00	39.87	40.44
Breast-cancer	9	24.11	24.92	<b>27.86</b>	27.17	26.83	25.89	22.22	23.34
Diabetes	8	25.23	27.36	<b>28.47*</b>	27.40	27.27	25.30	23.91	25.84
German	20	19.90	21.45	<b>24.83*</b>	23.90	22.48	20.31	18.70	19.65
Heart	13	31.69	31.11	<b>35.09*</b>	33.38	33.36	26.86	23.68	30.77
Titanic	3	28.43	28.11	<b>28.75</b>	28.55	28.45	28.01	27.62	27.89

\*) Statistically significantly better than all other methods for  $\alpha=0.05$

있다. 각 모델의 모수들은 cross-validation을 통해 결정되었다. SOM-G, OneSVM, SVDD와 비교했을 때, OneLVQ는 평균적으로 훨씬 낮은 FRR과 약간 높은 FAR을 보였다. 한편, SOM-L과 비교했을 때에는 거의 같은 FRR과 약간 낮은 FAR을 보였다. Parzen과 AANN은 실제 클래스 레이블에 관계없이 거의 모든 패턴을 이상으로 분류하였다. Parzen의 경우에 데이터가 고차원이면서 패턴의 수가 적을 때에는 분포 추정이 어렵다는 점이 잘 알려져 있다. AANN은 아주 적은 수의 데이터로 많은 수의 연결가중치를 추정해야 했기 때문에, overfitting을 피할 수 없었다. OneLVQ는 단 55개의 학습 데이터로 좋은 성능을 보여, 실제 문제에 대한 적용성을 나타내었다.

한 가지 강조할 점은 위의 이상 탐지 기법들 중에서, OneLVQ와 SVDD만이 학습 과정에서 이상 데이터를 활용한다는 점이다. 그러한 점만 제외하면, OneLVQ는 SOM 기반 기법들과 유사하고, SVDD는 OneSVM과 어느 정도 유사성을 보인다. OneLVQ와 SVDD가 각각 유사한 기법들보다 높은 성능을 보인다는 점을 고려하면, 학습 과정에서 이상 클래스의 정보를 활용하는 것이 이상 탐지 성능을 향상시킨다고 할 수 있다.

#### 4. 결론 및 토의

이 논문은 이상 탐지 문제에서 이상 데이터가 존재할 때 그것을 활용하는 방법에 초점을 두고 있다. 코드북 기반 기법의 이상 탐지 성능을 향상시키기 위하여 SOM-L과 OneLVQ 방법을 제안하였다. 우선, 두 가지 오류(FRR, FAR)의 trade-off를 고려하는 threshold 결정 방법을 제안하였다. 그리고, OneLVQ 방법에서는, 기존 LVQ의 오류 함수와 학습 규칙을 수정하여, 정상 데이터 뿐만 아니라 이상 데이터도 학습에 활용할 수 있도록 하였다.

2차원 합성 데이터로 실험한 결과, 정상 패턴만을 고려할 때 발생할 수 있는 문제점들을 OneLVQ

가 해결할 수 있었다. 또한, 6개의 벤치마크 데이터 셋과 타자 패턴 데이터 셋에 적용했을 때, 제안된 두 방법 모두 이상 탐지 기법으로서 경쟁력을 보였다. 특히, OneLVQ는 대부분의 문제에서 다른 널리 쓰이는 이상 탐지 기법들보다 좋은 성능을 나타내었다.

2005 한국경영과학회/대한산업공학회/한국계통통계학회/한국수리학회에 대한 FRR과 FAR의 평균값 (%)  
 2005년 5월 13일~14일, 충북대학교

Password	Parzen		OneLVQ		SOM-L		SOM-G		AANN		OneSVM		SVDD	
	FR	FA	FR	FA	FR	FA	FR	FA	FR	FA	FR	FA	FR	FA
	R	R	R	R	R	R	R	R	R	R	R	R	R	R
90200jdg	99.07	0.57	31.47	7.29	28.67	7.29	51.73	6.29	86.80	1.14	23.07	8.57	18.00	10.00
ahrfus88	89.87	0	8.00	2.29	7.47	2.14	19.07	0.14	68.00	0	11.47	0.29	10.13	0.71
anehwksu	92.67	0	12.93	8.29	11.20	12.43	24.13	5.14	68.67	1.86	15.73	7.29	14.80	8.14
autumnman	100	0	6.80	0	7.60	0	16.40	0	43.20	0	18.93	0	14.67	0
beaupowe	48.13	0	6.67	12.14	3.47	13.57	12.67	3.00	25.33	8.29	9.20	19.29	9.07	12.57
c.s.93/ksy	98.67	0	8.13	0	7.60	0	16.53	0	63.07	0	15.47	0	11.87	0
dhfpql.	85.73	0.29	9.73	12.71	10.13	13.29	18.27	8.00	70.40	2.00	13.20	9.14	13.73	8.57
dirdhfmw	97.60	0.14	6.27	11.71	5.07	13.86	18.93	5.86	45.60	5.43	12.27	9.57	11.87	8.71
dlfjs wp	98.40	0	12.00	1.71	12.67	2.86	24.27	1.14	64.27	0.57	22.27	1.00	19.07	1.57
dltjdgml	94.00	0	2.67	0	2.80	0.14	8.67	0	57.20	0	4.93	0.14	3.33	0.57
drizzle	89.87	0	7.87	9.29	9.07	9.29	23.33	2.71	39.87	2.14	14.80	2.89	13.47	3.57
dusru427	82.00	0	4.93	0	4.67	0.14	12.93	0	33.60	0	15.87	0.14	15.07	0
i love 3	94.93	0	6.27	12.00	5.20	14.14	10.40	9.71	49.47	4.71	7.20	11.43	6.00	12.14
love wjd	99.47	0	14.67	10.00	21.07	8.43	30.67	3.71	64.53	1.43	28.80	6.57	27.07	8.43
loveis.	82.53	0	9.07	1.14	8.53	6.14	16.53	0.29	58.93	0.43	13.47	0.49	13.20	0.89
manseiii	80.27	0	7.33	4.29	10.53	5.00	14.53	2.43	48.67	1.29	13.20	5.49	12.27	6.09
rhkdwo	76.40	0.49	12.67	6.14	13.20	10.14	30.80	3.14	76.93	0.49	18.40	9.19	17.07	9.49
rla sua	79.73	0	6.40	4.71	5.20	10.14	13.07	2.59	57.20	0.19	8.40	3.79	7.20	2.29
tjddmswjd	99.73	0	11.07	4.86	7.60	9.14	16.93	3.79	51.07	3.59	16.40	2.79	15.20	3.79
tmdwnsl1	87.73	0	9.87	5.29	11.33	5.14	28.40	0.79	58.13	0.49	15.73	2.89	14.80	3.19
yuhwalkk	96.00	0	4.80	0	5.73	0	12.00	0	47.07	0	11.60	0	8.00	0
Average	89.18	0.07	9.51	5.42	9.47	6.82	20.01	2.79	56.10	1.61	14.78	4.79	13.14	4.78

몇 가지 한계와 추후 과제를 지적할 수 있다. 첫째, 학습에 중요한 모수들인  $K$  와  $(C_1, C_2)$ 의 적당한 값을 결정하는 것이 어렵다. 현재로서는 cross-validation 등을 통하여 실험적으로 정할 수밖에 없는데, 한 쪽 클래스의 데이터가 부족한 상황에서 최선의 방법이라고는 할 수 없다. 둘째, 더 효율적인 초기화 방법이 필요하다. 셋째, OneLVQ가 효과적으로 작동하기 위해서 얼마나 많은 이상 패턴이 필요한지에 대한 연구도 수행되어야 한다.

**참고문헌**

[1] Bishop, C., “Novelty Detection and Neural Network Validation”, In *Proceedings of IEE Conference on Vision and Image Signal Processing* (1994), pp.217-222.

[2] Frosini, A., Gori, M., and Priami, P., “A Neural Network-based Model for Paper Currency Recognition and Verification”, *IEEE Transactions on Neural Networks*, Vol.7, No.6 (1996), pp.1482-1490.

[3] Gori M., Lastrucci, L., and Soda, G., “Autoassociator-based Models for Speaker Verification”, *Pattern Recognition Letters*, Vol.17 (1995), pp.241-250.

[4] Japkowicz, N., “Supervised versus Unsupervised Binary-Learning by Feed-forward Neural Networks”, *Machine Learning*, Vol.42, No.1-2 (2001), pp.97-122.

[5] Kohonen, T., *Self Organizing Maps*, Springer, Berlin, 2001.

[6] Lauer, M., “A Mixture Approach to Novelty Detection using Training Data with Outliers”, In *L. De Raedt, P. Flach (Eds), Proceedings of the 12th European Conference on Machine Learning* (2001), pp.300-311.

[7] Markou, M., and Singh, S., “Novelty Detection: A Review - Part 1: Statistical

Approaches”, *Signal Processing*, Vol. 83 (2003), pp.2481-2497.

[8] Markou, M., and Singh, S., “Novelty Detection: A Review - Part 2: Neural Network based Approaches”, *Signal Processing*, Vol.83, (2003), pp.2499-2521.

[9] Marsland, S., “Novelty Detection in Learning Systems”, *Neural Computing Surveys*, Vol.3 (2003), pp.157-195.

[10] Rätsch, G., Onoda, T., and Muller, K.R., “Soft margins for AdaBoost”, *Machine Learning*, Vol.42, No.3 (2001), pp.287-320.

[11] Schölkopf, B., Platt, J.C., Shawe-Taylor, J., Smola, A.J., and Williamson, R.C., “Estimating the Support of a High-Dimensional Distribution”, *Neural Computation*, Vol.13 (2001), pp.1443-1471.

[12] Tax, D.M.J., and Duin, R.P.W., “Support Vector Data Description”, *Machine Learning*, Vol.54 (2004), pp.45-66.

[13] Yu, E., and Cho, S., “Keystroke Dynamics Identity Verification - Its Problems and Practical Solutions”, *Computer and Security*, Vol.23, No.5 (2004), pp.428-440.